SUPPORTING INFORMATION

Selective small-molecule targeting of a triple helix encoded by the long non-coding RNA, *MALAT1*

Fardokht A. Abulwerdi[1,#], Wenbo Xu[2,3,#], Abeer A. Ageeli[4,#], Michael J. Yonkunas[4], Gayatri Arun[2], Hyeyeon Nam[5], John S. Schneekloth[6], Jr., Theodore Kwaku Dayie[5], David Spector[2], Nathan Baird[4*], and Stuart F. J. Le Grice[1*]

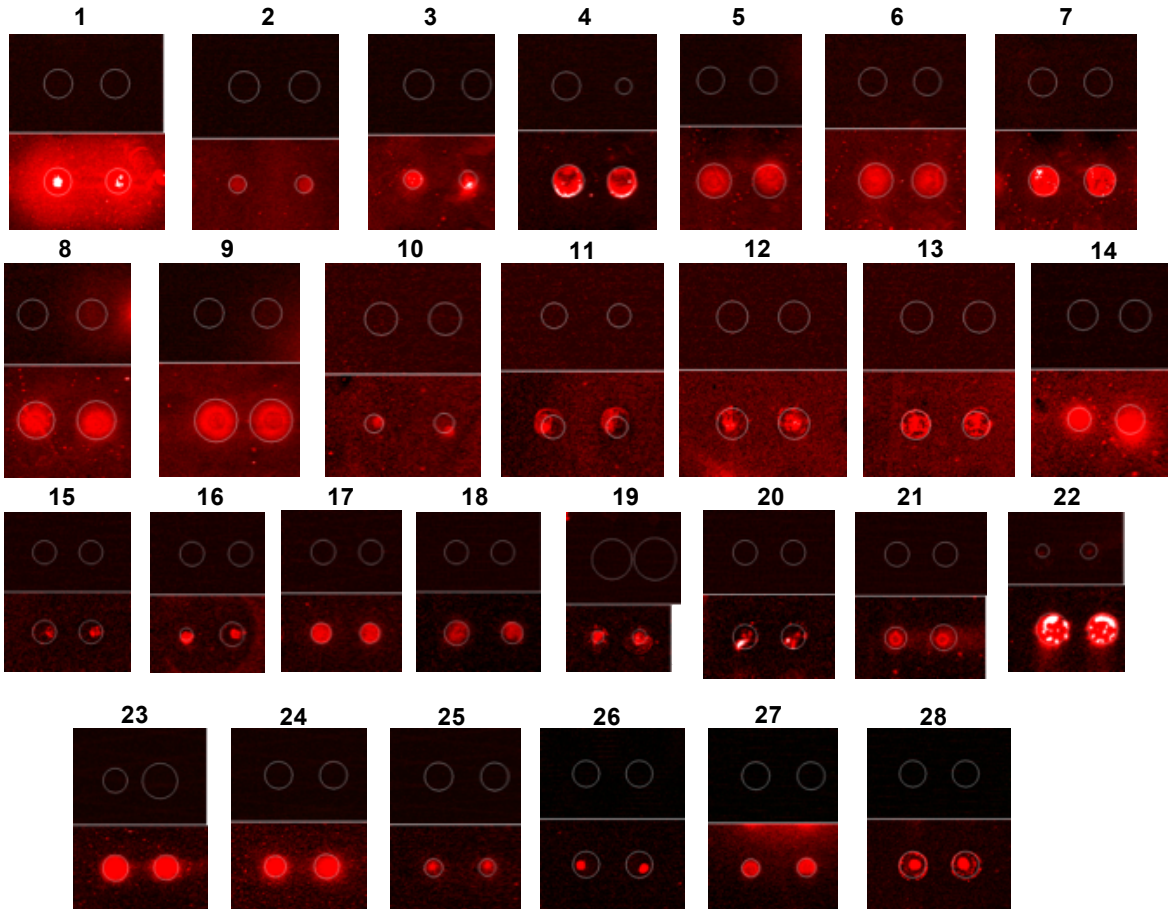

Figure S1. **Scanned images of hit candidates**. Each compound is printed in duplicate on the array. The top panel represents the array incubated with buffer while the bottom array represents the array incubated with 500 nM of AF647-mouse *Malat1* triple helix motif. Arrays are scanned at 635nm.

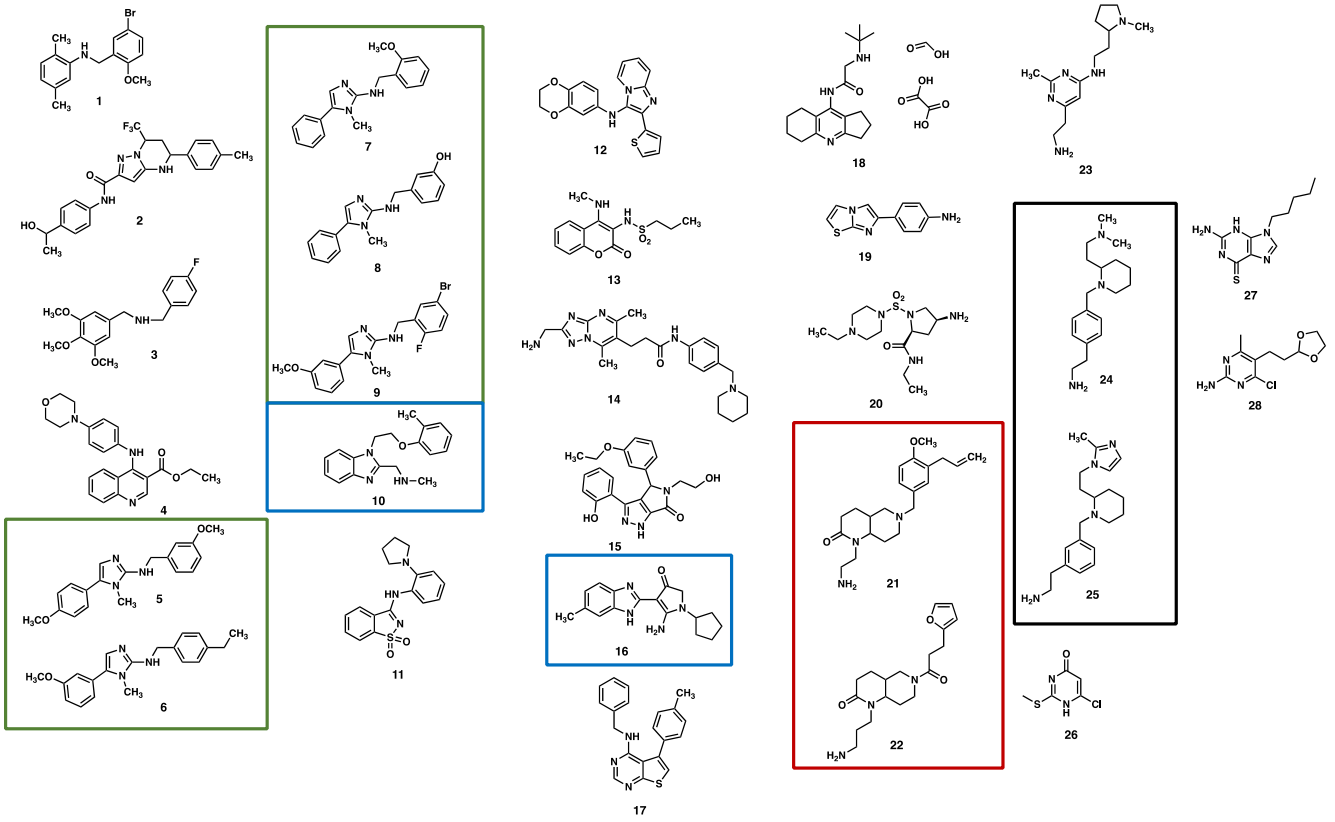

Figure S2. **Chemical structures of 28 hit molecules identified from SMM screening**. The boxed structures belong to a single class of compounds and include imidazoles (green box), benzimidazole (blue box), octahydronaphtharydinone (red box) and piperidine (black box).

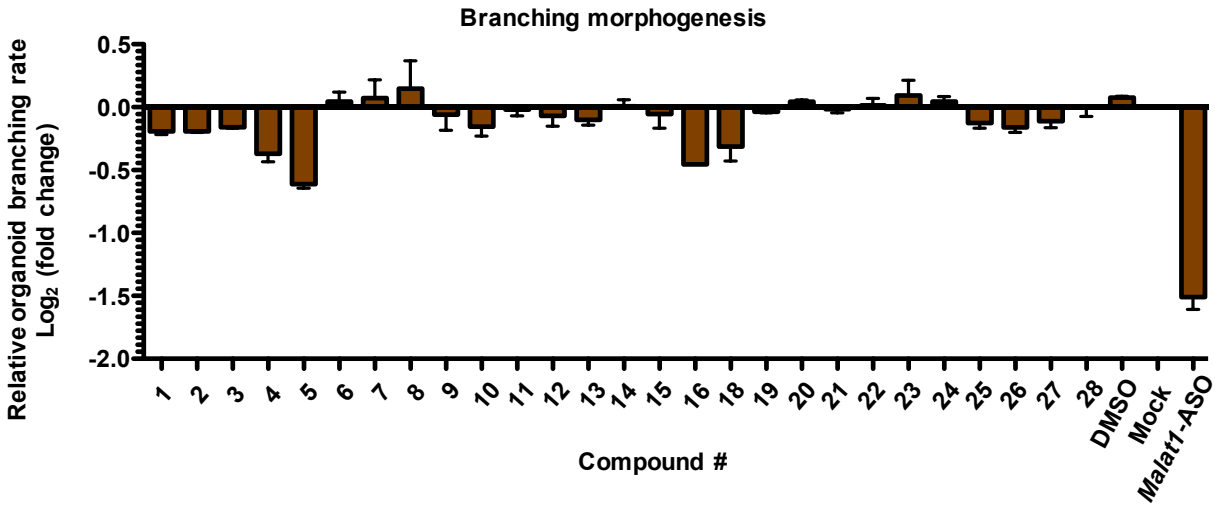

Figure S3. **Compounds 5 and 16 reduce organoid branching.** Relative organoid branching rate of MMTV-PyMT tumor organoids with treatments of Mock, DMSO, *Malat1*-ASO and compounds. n=3 biological replicates, bars represent +/- SEM.

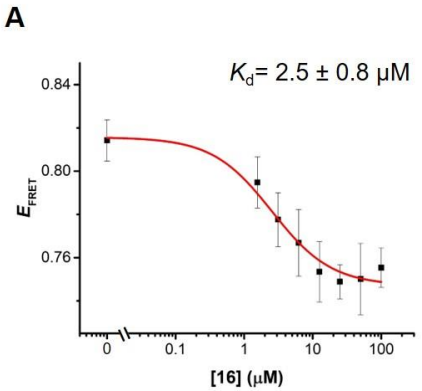

Figure S4. Compound 16 disrupts the M1^{ET} triplex. (A) Titration of compound **16** monitored by FRET using M1^{ET} in 1 mM $MgCl_2$ and 102.5 mM monovalent (equimolar NaCl and KCl). The K_d is 2.5 ± 0.8 µM. n=3 experiments, s.d.

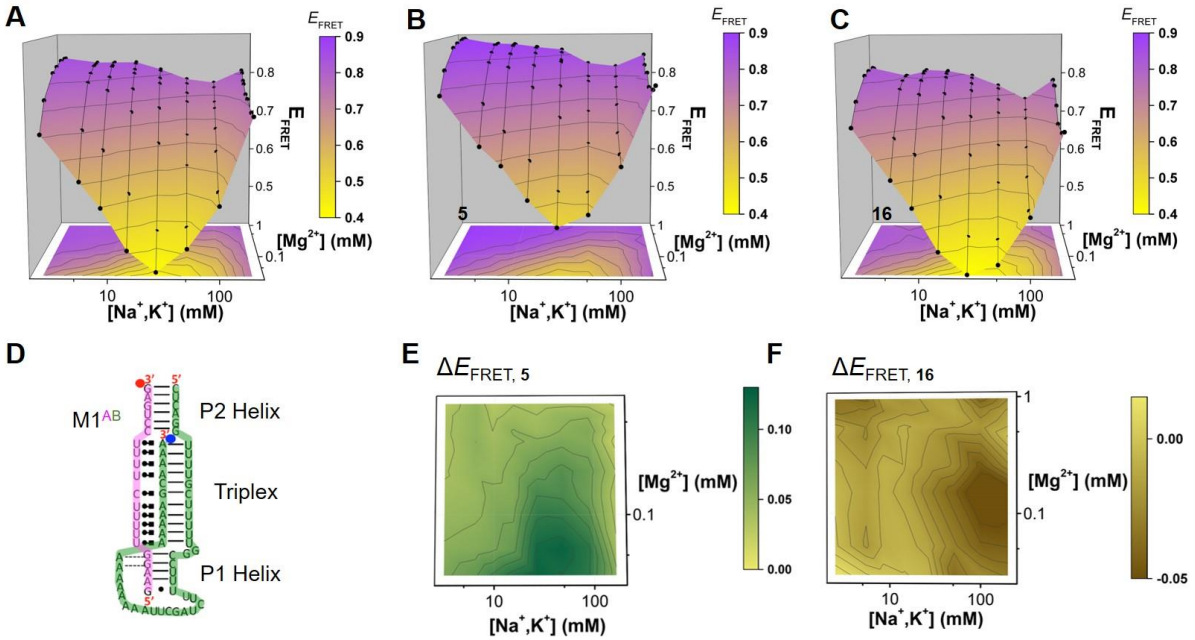

Figure S5. **Conformational landscapes of M1^{AB} confirm distinct interaction modes for compounds 5 and 16.** (**A**) Conformational landscape of M1^{AB} in the presence of vehicle (DMSO) alone reveals population of folded and disrupted triplex across a matrix of ionic conditions. Higher E_{FRET} represents more folded RNA and low E_{FRET} indicates less folded RNA. (**B-C**) Conformational landscapes of M1^{AB} in the presence of 10 µM compound **5** (**B**) and compound **16** (**C**), in which compound **5** increases E_{FRET} while compound **16** decreases E_{FRET} under all ionic conditions. In **A**-**C**, E_{FRET} was calculated as a function of 8 monovalent concentrations (equimolar of KCl and NaCl) and 8 magnesium concentrations. Experiments with DMSO and compound **5** were performed in duplicate. Experiments with compound **16** were performed once. Plotted here are single landscapes. (**D**) Secondary structure of M1^{AB} (M1^A in pink and M1^B in green) with the Cy3 and Cy5 positions highlighted by red and blue circles, respectively. (**E**) Difference plot of M1^{AB} E_{FRET} in the presence (10 µM) and absence (DMSO control) of compound **5**; n = 2 experiments. (**F**) Difference plot of M1^{AB} E_{FRET} in the presence (10 µM) and absence (DMSO control) of compound **16**; n = 1 experiment. In **E** and **F**, compound **5** increases E_{FRET} (green color) while compound **16** decreases the E_{FRET} (brown color) under most ionic conditions.

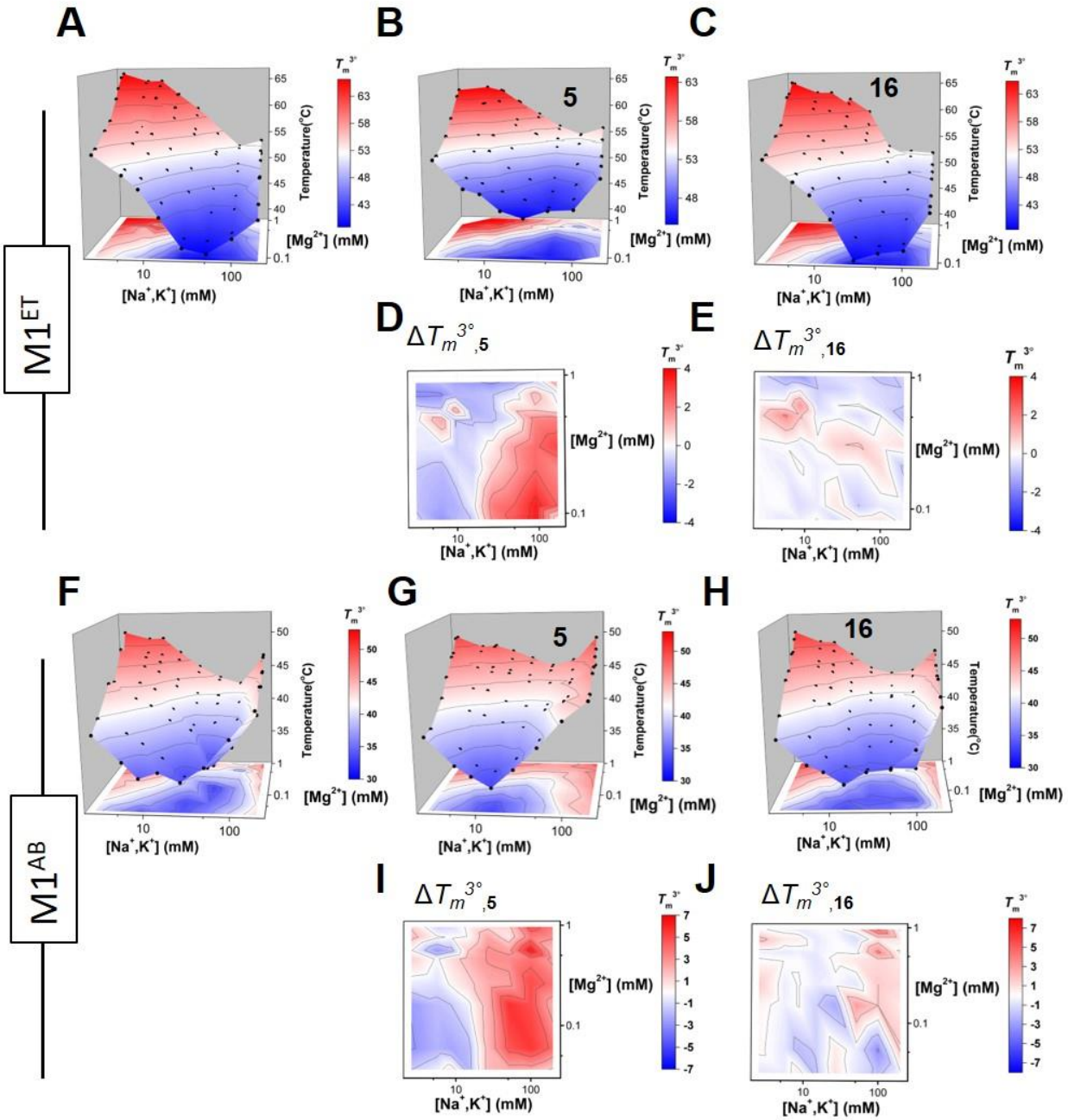

Figure S6. **Triplex thermal stability landscapes of M1^{ET} and M1^{AB} using differential scanning FRET (DS-FRET) confirm distinct interaction modes of compounds 5 and 16.** (**A**) Tertiary stability landscape of M1^{ET} in the absence of compound (DMSO control) across a matrix of ionic conditions. Thermal stabilities are colored on a gradient to indicate low (blue) and high (red) $T_m^{3°}$. (**B-C**) Stability landscapes in the presence of 10 µM compound **5** (**B**) or compound **16** (**C**). In **A**-**C**, tertiary melting temperature ($T_m^{3°}$) was calculated across a 8 x 8 matrix of monovalent concentrations from 2.56 - 202.56 mM (equimolar of KCl and NaCl) and magnesium concentrations from 0.1 - 1 mM. (**D-E**) Difference plots of $T_m^{3°}$ in the presence (10 µM) and absence (DMSO) of compound 5 (**D**) and compound **16** (**E**). Compound **5** slightly increases the stability of M1^{ET} under high monovalent concentrations but decreases triplex stability in low monovalent concentrations. Compound **16** shows no effect, within error, under all ionic conditions. (**F**) Tertiary stability landscape of M1^{AB} in the absence of compound (DMSO control) across a matrix of ionic conditions. Thermal stabilities are colored on a gradient to indicate low

(blue) and high (red) $T_m^{3^\circ}$. (**G-H**) Stability landscapes in the presence of 10 μM compound **5** (**G**) or compound **16** (**H**). In **F**-**H**, tertiary melting temperature ($T_m^{3^\circ}$) was calculated across a 8 x 8 matrix of monovalent concentrations from 2.56 - 202.56 mM (equimolar of KCl and NaCl) and magnesium concentrations from 0.1 - 1 mM. (**I-J**) Difference plots of $T_m^{3^\circ}$ in the presence (10 μM) and absence (DMSO control) of compound **5** (**I**) and compound **16** (**J**). As observed for $M1^{ET}$ in panels **D** and **E**, compound **5** slightly increases the stability of $M1^{AB}$ under high monovalent concentrations but decreases triplex stability in low monovalent concentrations. Compound **16** shows no effect, within error, under all ionic conditions.

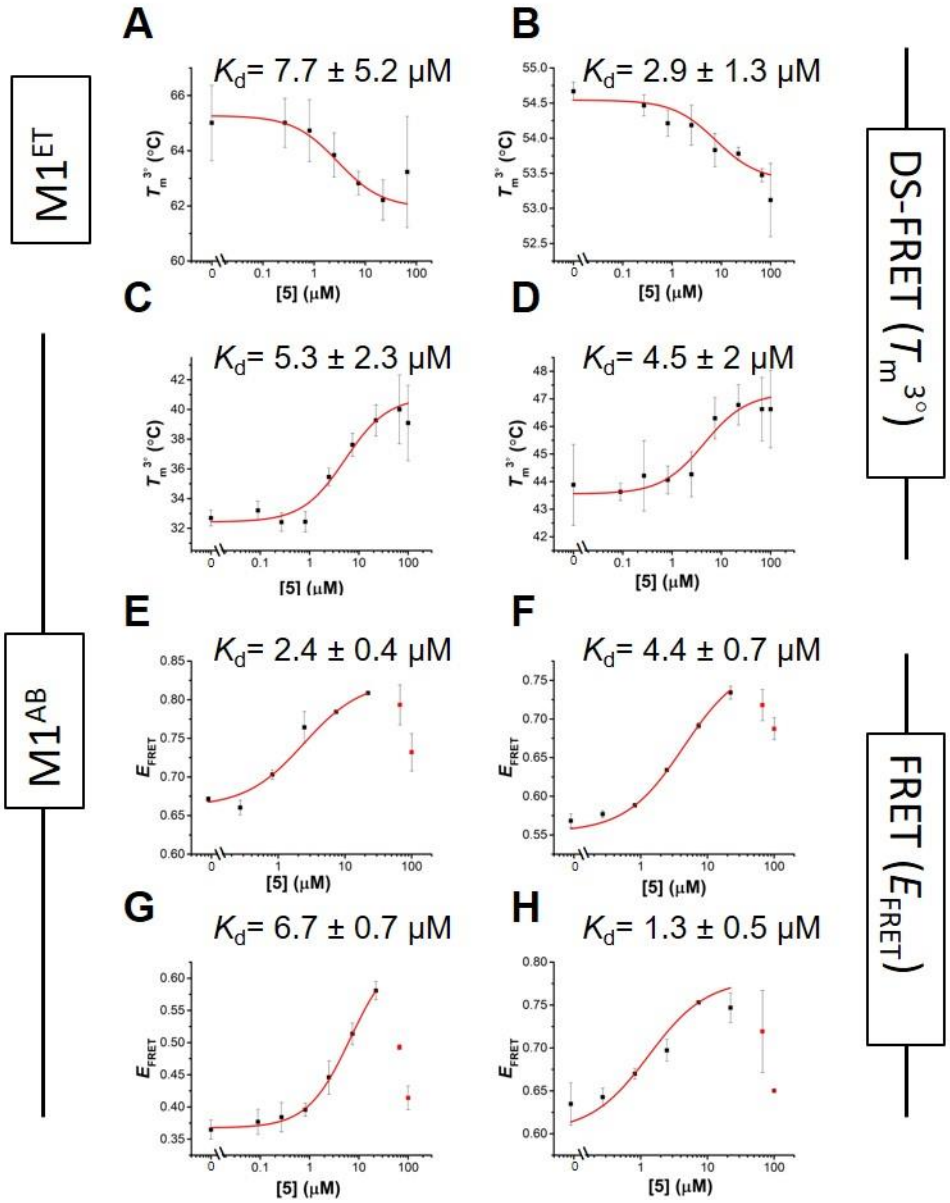

Figure S7. **Micromolar binding affinities (K_d) of compounds 5 under different salt conditions monitored by FRET and DS-FRET.** (**A**,**B**) Fits to $T_m{}^{3^\circ}$ titrations of compound **5** with M1^{ET} determine (**A**) K_d is 7.7 ± 5.2 μM in 0.1 mM $MgCl_2$ and 2.56 mM monovalent salt and (**B**) K_d is 2.9 ± 1.3 in 1 mM $MgCl_2$ and 2.56 mM monovalent salt. In panel **B**, only the Cy3 signal was plotted because FRET determination was precluded by Cy5 quenching at high compound concentrations. (**C**-**H**) K_d determination of compound **5** to M1^{AB} from titrations monitored by DS-FRET and FRET: (**C**) K_d is 5.3 ± 2.3 μM in 0.1 mM $MgCl_2$ and 52.56 mM monovalent (equimolar NaCl and KCl). (**D**) K_d is 4.5 ± 2 in 1 mM $MgCl_2$ and 102.56 mM monovalent (equimolar NaCl and KCl). (**E**) K_d is 2.4 ± 0.4 μM in 1 mM $MgCl_2$ and 2.56 mM monovalent (equimolar NaCl and KCl). (**F**) K_d is 4.4 ± 0.7 μM in 0.1 mM $MgCl_2$ and 2.56 mM monovalent salt. (**G**) K_d is 6.7 ± 0.7 μM in 0.1 mM $MgCl_2$ and 52.56 mM monovalent (equimolar NaCl and KCl). (**H**) K_d is 1.3 ± 0.5 μM in 1 mM $MgCl_2$ and 102.56 mM monovalent (equimolar NaCl and KCl). In panels **E** - **H** data were fit while masking the final two data points of the titration due to dye quenching at high concentrations of compound **5**. Experiments in all panels were performed in triplicate and error bars represent the standard deviation.

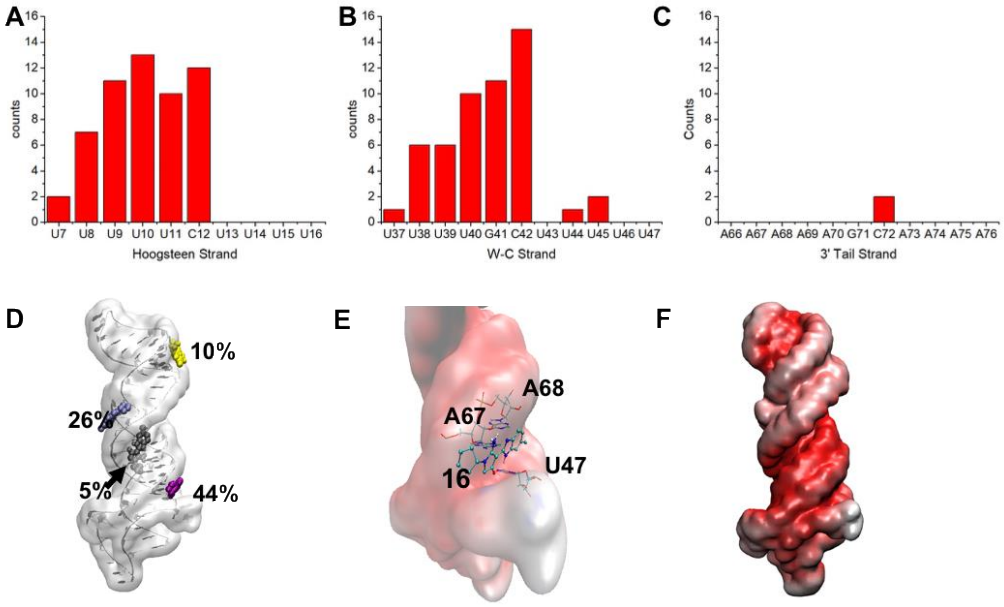

Figure S8. Histogram distribution analysis of interactions between the compound **5** single docking cluster with nucleotides in the **A.** Hoogsteen strand, **B.** Watson-Crick strand, and **C.** the 3' tail strand. A "count" is defined as any heavy atom or polar hydrogen within 3.5 Å of **5**. **D.** Alternate compound **16** surface binding clusters shown colored in a surface representation and labeled with their respective cluster populations. **E.** The "bulge cleft" formed by the 3' strand nucleotides A67 and A68 and U47 adjacent to bulge nucleotides G48 and G49. In addition to the slightly electronegative pocket, backbone sugar 2' hydroxyl and nucleobase carbonyl hydrogen bonding to the **16** amine make this pocket more probable over the lowest energy docking pose. **F.** A surface representation of the *MALAT1* ENE triplex crystal core with the electrostatic potential mapped to the surface using a blue, white, and red scale indicating positive, neutral, and negative charge respectively.

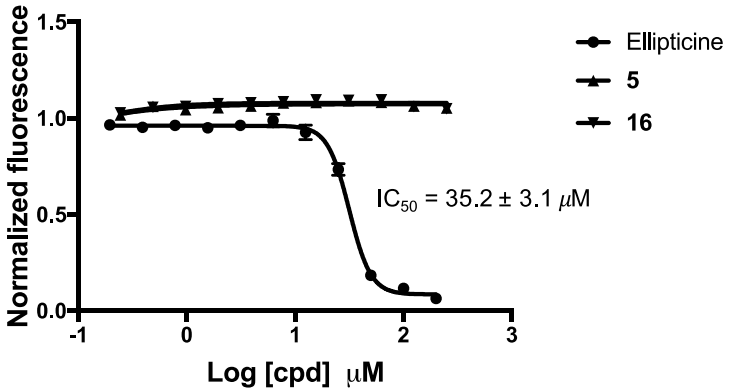

Figure S9. **5 and 16 do not intercalate the *MALAT1* ENE triple helix**. Dye displacement assay with a constant concentration of 94nt *MALAT1* triple helix construct and SYBRG II dye and in presence of increasing concentration of **5**, **16** and a known RNA intercalator, Ellipticine. Error bars represent the mean ± SEM.

Table S1. Binding affinities for compounds **5** and **16**

Experiment	RNA	Compound	K_d (μM)
FRET	$M1^{ET}$	**5**	2.3 ± 1.7[b]
		16	6.1 ± 1.8[c]
			2.5 ± 0.8[d]
	$M1^{AB}$	**5**	2.4 ± 0.4[a]
			4.4 ± 0.7[b]
			6.7 ± 0.7[c]
			1.3 ± 0.5[d]
DS-FRET	$M1^{ET}$	**5**	7.7 ± 5.2[b]
			2.9 ± 1.3[a*]
	$M1^{AB}$	**5**	5.3 ± 2.3[c]
			4.5 ± 2[d]
ITC	$M1^{TH}$	**5**	2.9 ± 1.6[e]
		16	6.1 ± 2.1[f]

[a]1 mM $MgCl_2$, 2.56 mM KCl
[b]0.1 mM $MgCl_2$, 2.56 mM KCl
[c]0.1 mM $MgCl_2$, 52.56 mM equimolar KCl and NaCl
[d]1 mM $MgCl_2$, 102.56 mM equimolar KCl and NaCl
[e]1 mM $MgCl_2$, 200 mM equimolar KCl and NaCl in 1% DMSO
[f]1 mM $MgCl_2$, 50 mM equimolar KCl and NaCl in 1%DMSO
*Cy3 signal is used for K_d determination instead of FRET due to Cy5 quenching at high compound concentrations.

Table S2. Thermodynamic binding parameters determined by ITC.

	5	**16**
K_d (μM)	2.9 ± 1.6	6.1 ± 2.1
n	0.8 ± 0.1	0.8 ± 0.1
ΔH (kJ mol^{-1})	-6.1 ± 0.5	-22.5 ± 5.4
ΔS (J mol^{-1} K^{-1})	86.2 ± 4.3	23.7 ± 15.5

Errors reported are the standard deviation from multiple independent experiments. n = 3 for compound **5**; n = 2 for compound **16**.

Table S3. The range of fit in the ITC parameters.

	K_d (μM)	**ΔH (kJ/mol)**	**ΔS (J/mol*K)**
Compound 5	2.9 — 3.9	-6.1 — -4.6	88.2 — 88.7
Compound 16	6.1 — 7.4	-22.5 — -18.9	23.5 — 34.5

The non-unity stoichiometry determined by ITC (see **Table S2**) results from imprecision in the calculated concentrations of the RNA or the compound. To evaluate the range of the fit parameters, we adjusted either the RNA or the compound concentration to achieve n = 1 stoichiometry. Adjustment to the RNA concentration does not change fit parameters, within error. Adjustment to the compound concentration impacts all of the fit parameters. Though the enthalpy and entropy are affected by this analysis, compound **5** binding remains entropically driven. The enthalpy of compound **16** binding is stronger than the enthalpy for compound **5** binding.